

KYP Lemma for Non-Strict Inequalities and the associated Minimax Theorem

A. Megretski, MIT (ameg@mit.edu)

August 17, 2010

Abstract

Several variations of the classical Kalman-Yakubovich-Popov Lemma, as well the associated minimax theorem are presented.

Notation and Terminology

\mathbb{Z}_+ is the set $\mathbb{Z}_+ = \{0, 1, 2, \dots\}$ of all non-negative integers. $j\mathbb{R} = \{s \in \mathbb{C} : \operatorname{Re}(s) = 0\}$, $\mathbb{C}_+ = \{s \in \mathbb{C} : \operatorname{Re}(s) > 0\}$, $\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$, and $\mathbb{D}_+ = \{z \in \mathbb{C} : |z| > 1\}$ are the frequently referenced subsets of the complex plane \mathbb{C} : the imaginary axis, the open right half plane, the unit circle, and the outside of the unit circle. $\mathbb{C}^{n,m} \supset \mathbb{R}^{n,m}$ are the sets of n -by- m matrices (complex and real), with the usual shortcuts $\mathbb{C}^n = \mathbb{C}^{n,1}$, $\mathbb{R}^n = \mathbb{R}^{n,1}$. For $M \in \mathbb{C}^{k,n}$, $M' \in \mathbb{C}^{n,k}$ is the Hermitian conjugate (the result of applying both transposition and complex conjugation to M), while $\bar{M} \in \mathbb{C}^{k,n}$ is the complex conjugate of M . For a real vector space V , V^\sharp is the real vector space of all linear functionals $f : V \mapsto \mathbb{R}$.

1 The Classical KYP Lemma

A number of alternative versions of the KYP Lemma, a classical result of the linear system theory, has been published over the last half century. The earlier formulations, such as [1], motivated by optimal linear feedback design applications, related positive definiteness (or semi-definiteness) of rational matrix-valued functions of a single complex variable on the real axis or on the unit circle (the so-called "frequency conditions") to the existence of "stabilizing" (or "marginally stabilizing") solutions of the associated Lur'e (algebraic Riccati) equations. Connections to dynamic programming and first order conditions of optimality allowed extensions to time-varying and distributed systems (see, for example, [2, 3]). Some of the more recent versions, such as [4], employ weaker assumptions to relate the frequency domain inequalities to feasibility of the semidefinite programs obtained by replacing the Lur'e or Riccati equations by the corresponding inequalities.

It appears that some useful versions of the KYP Lemma remain unpublished (or, at least, highly inaccessible). This paper aims to correct this by presenting several (assumedly) missing formulations.

1.1 KYP Lemma in Discrete Time

The classical KYP Lemma setup is defined by matrices $A \in \mathbb{C}^{n,n}$, $B \in \mathbb{C}^{n,m}$, $Q = Q' \in \mathbb{C}^{n+m,n+m}$. A and B are the coefficients of linear transformation

$$(x, u) \in \mathbb{C}^n \times \mathbb{C}^m \mapsto x_+ = Ax + Bu \in \mathbb{C}^n,$$

and Q is associated with the Hermitian form $\sigma : \mathbb{C}^n \times \mathbb{C}^m \mapsto \mathbb{R}$:

$$\sigma(x, u) = \begin{bmatrix} x \\ u \end{bmatrix}' Q \begin{bmatrix} x \\ u \end{bmatrix} \quad (x \in \mathbb{C}^n, u \in \mathbb{C}^m). \quad (1.1)$$

1.1.1 Stabilizing Completion of Squares in Discrete Time

This is one of the versions of the KYP Lemma, motivated by the linear quadratic optimal control design theory.

Theorem 1.1 *Assume that the pair (A, B) , where $A \in \mathbb{C}^{n,n}$ and $B \in \mathbb{C}^{n,m}$, is stabilizable, in the sense that there exists a matrix $K \in \mathbb{C}^{m,n}$ such that $zI_n - A - BK$ is invertible for all $z \in \mathbb{C}$, $|z| \geq 1$. Then for every matrix $Q = Q' \in \mathbb{C}^{n+m,n+m}$ (and σ defined in (1.1)) the following conditions are equivalent:*

(a) *there exist matrices $P = P' \in \mathbb{C}^{n,n}$, $C \in \mathbb{C}^{m,n}$, $D \in \mathbb{C}^{m,m}$ such that*

$$\sigma(x, u) + x'Px - (Ax + Bu)'P(Ax + Bu) = |Cx + Du|^2 \quad \forall x \in \mathbb{C}^n, u \in \mathbb{C}^m, \quad (1.2)$$

$$\det \begin{bmatrix} \lambda A - I_n & \lambda B \\ C & D \end{bmatrix} \neq 0 \quad \forall |\lambda| < 1; \quad (1.3)$$

(b) *the matrix*

$$\Pi(z) = \begin{bmatrix} (zI_n - A)^{-1}B \\ I_m \end{bmatrix}' Q \begin{bmatrix} (zI_n - A)^{-1}B \\ I_m \end{bmatrix}, \quad (1.4)$$

defined for $z \notin \Lambda(A) = \{z \in \mathbb{C} : \det(zI_n - A) = 0\}$, is positive definite for all $z \in \mathbb{T}$ except, possibly, a finite subset.

Moreover, when matrices A, B, Q in (b) are real, the corresponding matrices P, C, D from (a) can be chosen to be real as well.

A proof of Theorem 1.1 (as well as any other proof provided in this paper) can be found in the Appendix section.

We will refer to Theorem 1.1 as the "stabilizing completion of squares" version of the KYP Lemma, because the right side of (1.2) can be viewed of a "complete square" quadratic form, and (1.3) guarantees that the matrix $A - BD^{-1}C$ is well defined and "marginally stable" (has no eigenvalues z with $|z| > 1$).

1.1.2 Application: Optimal Program Control

The "stabilizing completion of squares" was originally motivated by an "abstract" optimal control question of finding the maximal lower bound of the functional

$$\Phi(x(\cdot), u(\cdot)) = \sum_{t=0}^{\infty} \sigma(x(t), u(t)) \rightarrow \inf \quad (1.5)$$

subject to linear equations

$$x(t+1) = Ax(t) + Bu(t), \quad x(0) = a, \quad (1.6)$$

and the "finite energy" constraint

$$\sum_{t=0}^{\infty} \{|x(t)|^2 + |u(t)|^2\} < \infty, \quad (1.7)$$

where A, B, Q, a are fixed, and $x : \mathbb{Z}_+ \mapsto \mathbb{C}^n$, $u : \mathbb{Z}_+ \mapsto \mathbb{C}^m$ are infinite dimensional decision variables. The following statement, which follows directly from Theorem 1.1, explains the relation between the optimization setup (1.5)-(1.7) and Theorem 1.1.

Theorem 1.2 *If the stabilizability assumption as well as conditions (a),(b) from Theorem 1.1 are satisfied then the infimum in (1.5)-(1.7) equals $-a'Pa$, and the sum*

$$\sum_{t=0}^{\infty} |Cx_i(t) + Du_i(t)|^2$$

converges to zero if and only if $\Phi(x_i(\cdot), u_i(\cdot))$ converges to $-a'Pa$.

1.1.3 Strict Linear Matrix Inequalities

In many applications, the "stabilization" constraint is irrelevant, which motivates the following "strict linear matrix inequality (LMI)" version of the KYP Lemma.

Theorem 1.3 *For arbitrary matrices $A \in \mathbb{C}^{n,n}$, $B \in \mathbb{C}^{n,m}$, $Q = Q' \in \mathbb{C}^{n+m,n+m}$ the following conditions are equivalent:*

- (a) *there exists $P = P' \in \mathbb{C}^{n,n}$ such that the Hermitian form $\sigma_P : \mathbb{C}^n \times \mathbb{C}^m \mapsto \mathbb{R}$ defined by*

$$\sigma_P(x, u) = \sigma(x, u) + x'Px - (Ax + Bu)'P(Ax + Bu) \quad (1.8)$$

is positive definite;

- (b) *the Hermitian form σ is positive definite on the subspace*

$$\mathcal{L}(z) = \{(x, u) \in \mathbb{C}^n \times \mathbb{C}^m : zx = Ax + Bu\} \quad (1.9)$$

for all $z \in \mathbb{T}$.

Moreover, when matrices A, B, Q in (b) are real, the corresponding matrix P from (a) can be chosen to be real as well.

1.1.4 Non-Strict Linear Matrix Inequalities

Since $\sigma_P(x, u) = \sigma(x, u)$ for $(x, u) \in \mathcal{L}(z)$, $z \in \mathbb{T}$, existence of a $P = P' \in \mathbb{C}^{n,n}$ for which the Hermitian form (1.8) is positive semidefinite implies that σ is positive semidefinite on $\mathcal{L}(z)$ for all $z \in \mathbb{T}$. In general, the inverse implication is not true: for example, when $A = 0$, $B = 0$, and $\sigma(x, u) = \operatorname{Re}(x'u)$, the subspace $\mathcal{L}(z)$, for all $z \in \mathbb{T}$, consists of all pairs $(0, u)$ with $u \in \mathbb{C}$, and, accordingly, $\sigma(x, u) = 0$ for $(x, u) \in \mathcal{L}(z)$, $z \in \mathbb{T}$. However, there exists no $P = P' \in \mathbb{C}^{1,1}$ (i.e. $P \in \mathbb{R}$) for which $\sigma_P(x, u) = \operatorname{Re}(x'u) + P|x|^2$ is positive semidefinite.

The following statement is a "non-strict LMI" version of the KYP Lemma which trades strict positivity for controllability of the pair (A, B) . Recall that a pair (A, B) of matrices $A \in \mathbb{C}^{n,n}$, $B \in \mathbb{C}^{n,m}$ is called *controllable* when the matrix $[\lambda I_n - A, B]$ is right invertible for all $\lambda \in \mathbb{C}$.

Theorem 1.4 *Assume that the pair (A, B) of matrices $A \in \mathbb{C}^{n,n}$, $B \in \mathbb{C}^{n,m}$ is controllable. Then for every matrix $Q = Q' \in \mathbb{C}^{n+m, n+m}$ the following conditions are equivalent:*

- (a) *there exists $P = P' \in \mathbb{C}^{n,n}$ such that the Hermitian form $\sigma_P : \mathbb{C}^n \times \mathbb{C}^m \mapsto \mathbb{R}$ defined by (1.8) is positive semidefinite;*
- (b) *the Hermitian form σ is positive semidefinite on the subspace $\mathcal{L}(z)$ defined by (1.9) for all $z \in \mathbb{T}$.*

Moreover, when matrices A, B, Q in (b) are real, the corresponding matrix P from (a) can be chosen to be real as well.

1.2 KYP Lemma in Continuous Time

Continuous time (CT) versions of the KYP lemma are similar to their DT counterparts.

Theorem 1.5 *Assume that the pair (A, B) , where $A \in \mathbb{C}^{n,n}$ and $B \in \mathbb{C}^{n,m}$, is stabilizable, in the sense that there exists a matrix $K \in \mathbb{C}^{m,n}$ such that $sI_n - A - BK$ is invertible for all $s \in \mathbb{C}$ with $\operatorname{Re}(s) \geq 0$. Then for every matrix $Q = Q' \in \mathbb{C}^{n+m, n+m}$ the following conditions are equivalent:*

- (a) *there exist matrices $P = P' \in \mathbb{C}^{n,n}$, $C \in \mathbb{C}^{m,n}$, $D \in \mathbb{C}^{m,m}$ such that*

$$\sigma(x, u) - 2\operatorname{Re}[x'P(Ax + Bu)] = |Cx + Du|^2 \quad \forall x \in \mathbb{C}^n, u \in \mathbb{C}^m, \quad (1.10)$$

$$\det \begin{bmatrix} A - sI_n & B \\ C & D \end{bmatrix} \neq 0 \quad \forall s \in \mathbb{C}_+; \quad (1.11)$$

- (b) *the matrix $\Pi(s)$ defined by (1.4) is positive definite for all $s \in j\mathbb{R}$ except, possibly, a finite subset.*

Moreover, when matrices A, B, Q in (b) are real, the corresponding matrices P, C, D from (a) can be chosen to be real as well.

Theorem 1.6 For arbitrary matrices $A \in \mathbb{C}^{n,n}$, $B \in \mathbb{C}^{n,m}$, $Q = Q' \in \mathbb{C}^{n+m,n+m}$ the following conditions are equivalent:

- (a) there exists $P = P' \in \mathbb{C}^{n,n}$ such that the Hermitian form $\sigma_P : \mathbb{C}^n \times \mathbb{C}^m \mapsto \mathbb{R}$ defined by

$$\sigma_P(x, u) = \sigma(x, u) + 2\operatorname{Re}[x'P(Ax + Bu)] \quad (1.12)$$

is positive definite;

- (b) the Hermitian form σ is positive definite on the subspace $\mathcal{L}(s)$ for all $s \in j\mathbb{R} \cup \{\infty\}$, where $\mathcal{L}(z)$ is defined by (1.9) for $z \in \mathbb{C}$, and

$$\mathcal{L}(\infty) = \{0\} \times \mathbb{C}^m. \quad (1.13)$$

Moreover, when matrices A, B, Q in (b) are real, the corresponding matrix P from (a) can be chosen to be real as well.

Theorem 1.7 Assume that the pair (A, B) of matrices $A \in \mathbb{C}^{n,n}$, $B \in \mathbb{C}^{n,m}$ is controllable. Then for every matrix $Q = Q' \in \mathbb{C}^{n+m,n+m}$ the following conditions are equivalent:

- (a) there exists $P = P' \in \mathbb{C}^{n,n}$ such that the Hermitian form $\sigma_P : \mathbb{C}^n \times \mathbb{C}^m \mapsto \mathbb{R}$ defined by (1.12) is positive semidefinite;
- (b) the Hermitian form σ is positive semidefinite on the subspace $\mathcal{L}(s)$ defined by (1.9) for all $s \in j\mathbb{R}$.

Moreover, when matrices A, B, Q in (b) are real, the corresponding matrix P from (a) can be chosen to be real as well.

2 A Minimax Theorem

It is easy to show that the inequality

$$\inf_{v \in V} \sup_{w \in W} g(v, w) \geq \sup_{w \in W} \inf_{v \in V} g(v, w) \quad (2.14)$$

holds for arbitrary sets V, W and arbitrary real-valued function $g : V \times W \mapsto \mathbb{R}$. The term *minimax theorem* refers to a family of statements providing conditions (usually involving convexity of g with respect to v and concavity of g with respect to w) under which the inequality in (2.14) is actually an equality, i.e.

$$\inf_v \sup_w g(v, w) = \sup_w \inf_v g(v, w). \quad (2.15)$$

In this section, we are particularly interested in a specific minimax statement partially motivated by the KYP Lemma.

2.1 Minimax Theorems for Discrete Time LTI Systems

For a positive integer m let ℓ_m^2 denote the standard real Hilbert space of all one-sided real m -vector valued square summable sequences, i.e. functions $u : \mathbb{Z}_+ \mapsto \mathbb{R}^m$ such that

$$\|u(\cdot)\|^2 = \sum_{t=0}^{\infty} |u(t)|^2 < \infty. \quad (2.16)$$

Given a Schur matrix $A \in \mathbb{R}^{n,n}$ (i.e. such that $zI_n - A$ is not singular for $|z| \geq 1$), a vector $a \in \mathbb{R}^n$, and matrices $B_1 \in \mathbb{R}^{n,k}$, $B_2 \in \mathbb{R}^{n,q}$, $Q \in \mathbb{R}^{n+k+q,n+k+q}$, consider the functional $g : \ell_k^2 \times \ell_q^2 \mapsto \mathbb{R}$ defined by

$$g(v(\cdot), w(\cdot)) = \sum_{t=0}^{\infty} \sigma(x(t), v(t), w(t)) : \quad x(t+1) = Ax(t) + B_1v(t) + B_2w(t), \quad x(0) = a, \quad (2.17)$$

where

$$\sigma(x, v, w) = \begin{bmatrix} x \\ v \\ w \end{bmatrix}' Q \begin{bmatrix} x \\ v \\ w \end{bmatrix} \quad (x \in \mathbb{R}^n, v \in \mathbb{R}^k, w \in \mathbb{R}^q). \quad (2.18)$$

Consider also the associated matrix $\Pi = \Pi(z)$ defined by (1.4) with $B = [B_1, B_2]$, and its partition

$$\Pi(z) = \begin{bmatrix} \Pi_{11}(z) & \Pi_{12}(z) \\ \Pi_{21}(z) & \Pi_{22}(z) \end{bmatrix}, \quad \Pi_{11}(z) \in \mathbb{C}^{k,k}, \quad \Pi_{22}(z) \in \mathbb{C}^{q,q}. \quad (2.19)$$

Our objective is to formulate conditions, in terms of matrices Π_{ij} , which guarantee that equality (2.15) is satisfied for all $a \in \mathbb{R}^n$ for the functional $g : \ell_k^2 \times \ell_q^2 \mapsto \mathbb{R}$ defined by (2.17). We are also interested in formulating conditions which ensure that the associated partial infimum and supremum

$$g_v(w) = \inf_v g(v, w), \quad g_w(v) = \sup_w g(v, w) \quad (2.20)$$

are finite, and that the resulting functions $g_v : \ell_q^2 \mapsto \mathbb{R}$, $g_w : \ell_k^2 \mapsto \mathbb{R}$ are continuous in the standard Hilbert space metrics of ℓ_q^2 and ℓ_k^2 , respectively.

2.1.1 A Counterexample

The Parseval identity can be used to show that $g(v, w)$ from (2.17) is convex with respect to v if and only if $\Pi_{11}(z) \geq 0$ for all $z \in \mathbb{T}$. Similarly, $g(v, w)$ is concave with respect to w if and only if $\Pi_{22}(z) \leq 0$ for all $z \in \mathbb{T}$. However, these assumptions are far from being

sufficient to assure that the minimax identity (2.15) is satisfied, as demonstrated by the example with

$$A = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad B_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad B_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad a = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad Q = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & -1 & 1 & 1 \\ -1 & 1 & 1 & -1 \\ 0 & 1 & -1 & -1 \end{bmatrix},$$

i.e. when

$$g(v(\cdot), w(\cdot)) = \sum_{t=0}^{\infty} \{|v(t) - x_1(t)|^2 - 2v(t)[w(t) - x_2(t)] - |w(t) - x_2(t)|^2\},$$

$$\text{subject to } x_1(t+1) = v(t), \quad x_2(t+1) = w(t), \quad x_1(0) = 1, \quad x_2(0) = 0,$$

and

$$\Pi(z) = \begin{bmatrix} |z-1|^2 & z-1 \\ z'-1 & -|z-1|^2 \end{bmatrix}.$$

Using the fact that the set of all possible sequences $w - x_2$ is dense in $V = W = \ell^2$, we conclude that

$$\sup_{w \in W} g(v, w) = \sum_{t=0}^{\infty} \{|v(t) - x_1(t)|^2 + |v(t)|^2\} \geq |v(0) - 1|^2 + |v(0)|^2 \geq 0.5,$$

and hence the left side in (2.15) is not smaller than 0.5.

On the other hand, using the fact that

$$\sum_{t=0}^{\infty} \{x_1(t)x_2(t) - v(t)w(t)\} = \sum_{t=0}^{\infty} \{x_1(t)x_2(t) - x_1(t+1)x_2(t+1)\} = x_1(0)x_2(0) = 0,$$

one can re-write the sum for g as

$$g(v(\cdot), w(\cdot)) = \sum_{t=0}^{\infty} \{|v(t) - x_1(t)|^2 + 2[v(t) - x_1(t)]x_2(t) - |w(t) - x_2(t)|^2\}.$$

Since the set of all possible sequences $v - x_1$ is dense in ℓ^2 , we conclude that

$$\inf_{v \in V} g(v, w) = \sum_{t=0}^{\infty} \{-|x_2(t)|^2 - |w(t) - x_2(t)|^2\},$$

and hence the right side in (2.15) is zero: the minimax equality does not hold in this case.

It is instructive to note that, in this case, both functions g_v and g_w from (2.20) are finite and continuous in the standard Hilbert space topology of ℓ^2 , but the minimax identity is still not valid.

2.1.2 A Sufficient Condition of Minimax

The following statement shows that the minimax identity (2.15) holds for the functional (2.17) when there exist $\epsilon > 0$ and $z_0 \in \mathbb{T}$ such that

$$\begin{bmatrix} \Pi_{11}(z) & \epsilon \Pi_{12}(z) \\ \epsilon \Pi_{21}(z) & -\Pi_{22}(z) \end{bmatrix} \geq 0 \quad \forall z \in \mathbb{T}, \quad \Pi_{11}(z_0) > 0, \quad \Pi_{22}(z_0) < 0. \quad (2.21)$$

Theorem 2.1 *Let $A \in \mathbb{R}^{n,n}$ be a Schur matrix. Assume that matrices $B_1 \in \mathbb{R}^{n,k}$, $B_2 \in \mathbb{R}^{n,q}$, $Q \in \mathbb{R}^{n+k+q,n+k+q}$ are such that condition (2.21), where Π_{ij} are defined by (2.19) and (1.4) with $B = [B_1, B_2]$, is satisfied for some $\epsilon > 0$ and $z_0 \in \mathbb{T}$. Then for every $a \in \mathbb{R}^n$ for the functional $g : \ell_k^2 \times \ell_q^2 \mapsto \mathbb{R}$ defined by (2.17), (2.18)*

- (a) *the partial optimal values in (2.20) are continuous in the standard norm topologies of ℓ_q^2 and ℓ_k^2 ;*
- (b) *the minimax identity in (2.15) is satisfied.*

2.1.3 Minimax and Integral Quadratic Constraints

For a positive integer m let ℓ_m denote the set of all functions $u : \mathbb{Z}_+ \mapsto \mathbb{R}^m$ (in particular, ℓ_m^2 is a subset of ℓ_m). In modeling discrete time dynamical systems, m -dimensional *signals* can be represented by the elements of ℓ_m . Accordingly, a DT system Δ with k -dimensional input v and q -dimensional output w is viewed as a subset $\Delta \subset \ell_k \times \ell_q$. Let us call such system $\Delta \subset \ell_k \times \ell_q$ *weakly causally stable* if for every $T \in \mathbb{Z}_+$, $(v, w) \in \Delta$, and $v_* \in \ell_k^2$ such that $v(t) = v_*(t)$ for all $t \leq T$ there exists a sequence of elements $(v_i, w_i) \in \Delta \cap (\ell_k^2 \times \ell_q^2)$, such that $v_i(t) = v(t)$ and $w_i(t) = w(t)$ for all $t \leq T$, and $\|v_i - v_*\| \rightarrow 0$ as $i \rightarrow \infty$.

Given real matrices $A \in \mathbb{R}^{n,n}$, $B_1 \in \mathbb{R}^{n,k}$, $B_2 \in \mathbb{R}^{n,q}$, $Q \in \mathbb{R}^{n+k+q,n+k+q}$, where A is a Schur matrix, and a subset $X_0 \subset \mathbb{R}^n$, let us say that system $\Delta \subset \ell_k \times \ell_q$ satisfies the *conditional* Integral Quadratic Constraint (IQC) defined by A , B_1 , B_2 , Q , X_0 if there exists a continuous function $\kappa : \mathbb{R}^n \times \mathbb{R}^k \times \mathbb{R}^q \mapsto \mathbb{R}$ such that

$$\sum_{t=0}^{\infty} \sigma(x(t), v(t), w(t)) \geq -\kappa(x_0, v(0), w(0)) \quad (2.22)$$

for all $(v, w) \in \Delta \cap (\ell_k^2 \times \ell_q^2)$, $x_0 \in X_0$, where $x(\cdot)$ is defined by $v(\cdot)$, $w(\cdot)$, and x_0 according to

$$x(t+1) = Ax(t) + B_1 v(t) + B_2 w(t), \quad x(0) = x_0. \quad (2.23)$$

Similarly, let us say that $\Delta \subset \ell_k \times \ell_q$ satisfies the *complete* IQC defined by A , B_1 , B_2 , Q , X_0 if there exists a continuous function $\kappa : \mathbb{R}^n \times \mathbb{R}^k \times \mathbb{R}^q \mapsto \mathbb{R}$ such that

$$\sum_{t=0}^T \sigma(x(t), v(t), w(t)) \geq -\kappa(x_0, v(0), w(0)) \quad (T \geq 0) \quad (2.24)$$

for all $(v, w) \in \Delta$, $x_0 \in X_0$, and $x \in \ell_n$ satisfying (2.23).

An important step in the IQC framework of nonlinear system analysis is to establish that a particular conditional IQC (2.22) implies the corresponding complete IQC (2.24). The implication is not always true: for example, when

$$\begin{aligned}\Delta &= \{(v, w) \in \ell \times \ell : w(t+1) = v(t) \ \forall t \in \mathbb{Z}_+\}, \\ A &= B_1 = B_2 = 0, \quad X_0 = \{0\}, \quad \sigma(x, v, w) = |w|^2 - |v|^2\end{aligned}$$

then the conditional IQC (2.22) is satisfied with $\kappa(x_0, v_0, w_0) = |v_0|^2$, but the associated complete IQC (2.24) does not take place for any function κ .

The following statement, based on the minimax identity established in Theorem 2.1, provides sufficient conditions, expressed in terms of matrices A , B , and Q , under which the conditional IQC from (2.22) implies the complete IQC from (2.24).

Theorem 2.2 *Let $\Delta \subset \ell_k \times \ell_q$ be a weakly causally stable system which satisfies the conditional IQC defined by real matrices $A \in \mathbb{R}^{n,n}$, $B_1 \in \mathbb{R}^{n,k}$, $B_2 \in \mathbb{R}^{n,q}$, $Q \in \mathbb{R}^{n+k+q, n+k+q}$, where A is a Schur matrix, and a subset $X_0 \subset \mathbb{R}^n$. Assume that*

- (a) *condition (2.21), where Π_{ij} are defined by (2.19) and (1.4) with $B = [B_1, B_2]$, is satisfied for some $\epsilon > 0$ and $z_0 \in \mathbb{T}$;*
- (b) *there exist real matrices $C \in \mathbb{R}^{k,n}$, $D_1 \in \mathbb{R}^{k,k}$, $D_2 \in \mathbb{R}^{k,q}$ such that the quadratic form σ defined by (2.18) satisfies the inequality*

$$\sigma(x, v, w) \leq |Cx + D_1v + D_2w|^2 \quad \forall x \in \mathbb{R}^n, v \in \mathbb{R}^k, w \in \mathbb{R}^q, \quad (2.25)$$

and

$$\det \begin{bmatrix} \lambda A - I & \lambda B_1 \\ C & D_1 \end{bmatrix} \neq 0 \quad \forall \lambda \in \mathbb{C}, |\lambda| < 1. \quad (2.26)$$

Then Δ satisfies the complete IQC defined by A, B_1, B_2, Q, X_0 .

2.2 Minimax Theorems for Continuous Time LTI Systems

For a positive integer m let L_m^2 denote the standard real Hilbert space of all real m -vector valued square integrable functions $u : [0, \infty) \mapsto \mathbb{R}^m$, equipped with the norm

$$\|u(\cdot)\|^2 = \int_0^\infty |u(t)|^2 dt < \infty. \quad (2.27)$$

Given a Hurwitz matrix $A \in \mathbb{R}^{n,n}$ (i.e. such that $sI_n - A$ is not singular for $\text{Re}(s) \geq 0$), a vector $a \in \mathbb{R}^n$, and matrices $B_1 \in \mathbb{R}^{n,k}$, $B_2 \in \mathbb{R}^{n,q}$, $Q \in \mathbb{R}^{n+k+q, n+k+q}$, consider the functional $g : L_k^2 \times L_q^2 \mapsto \mathbb{R}$ defined by

$$g(v(\cdot), w(\cdot)) = \int_0^\infty \sigma(x(t), v(t), w(t)) dt : \dot{x}(t) = Ax(t) + B_1v(t) + B_2w(t), \quad x(0) = a, \quad (2.28)$$

where $\sigma(\cdot)$ is defined by (2.18). Consider also the associated matrix $\Pi(\cdot)$ defined by (1.4) with $B = [B_1, B_2]$, and its partition (2.19).

Our objective is to formulate conditions, in terms of matrices Π_{ij} , which guarantee that equality (2.15) is satisfied for all $a \in \mathbb{R}^n$ for the functional $g : L_k^2 \times L_q^2 \mapsto \mathbb{R}$ defined by (2.28). We are also interested in formulating conditions which ensure that the associated partial infimum and supremum (2.20) are finite, and that the resulting functions $g_v : L_q^2 \mapsto \mathbb{R}$, $g_w : L_k^2 \mapsto \mathbb{R}$ are continuous in the standard Hilbert space metrics of L_q^2 and L_k^2 , respectively.

2.2.1 A Sufficient Condition of Minimax

The following statement shows that the minimax identity (2.15) holds for the functional (2.28) when there exist $\epsilon > 0$ and $s_0 \in j\mathbb{R}$ such that

$$\begin{bmatrix} \Pi_{11}(s) & \epsilon \Pi_{12}(s) \\ \epsilon \Pi_{21}(s) & -\Pi_{22}(s) \end{bmatrix} \geq 0 \quad \forall s \in j\mathbb{R}, \quad \Pi_{11}(s_0) > 0, \quad \Pi_{22}(s_0) < 0. \quad (2.29)$$

Theorem 2.3 *Let $A \in \mathbb{R}^{n,n}$ be a Hurwitz matrix. Assume that matrices $B_1 \in \mathbb{R}^{n,k}$, $B_2 \in \mathbb{R}^{n,q}$, $Q \in \mathbb{R}^{n+k+q,n+k+q}$ are such that condition (2.29), where Π_{ij} are defined by (2.19) and (1.4) with $B = [B_1, B_2]$, is satisfied for some $\epsilon > 0$ and $s_0 \in j\mathbb{R}$. Then for every $a \in \mathbb{R}^n$ for the functional $g : L_k^2 \times L_q^2 \mapsto \mathbb{R}$ defined by (2.28), (2.18)*

- (a) *the partial optimal values in (2.20) are continuous in the standard norm topologies of L_q^2 and L_k^2 ;*
- (b) *the minimax identity in (2.15) is satisfied.*

2.2.2 Minimax and Continuous Time IQC

For a positive integer m let L_m denote the set of all locally square integrable functions $u : [0, \infty) \mapsto \mathbb{R}^m$ (in particular, L_m^2 is a subset of L_m). In modeling continuous time dynamical systems, m -dimensional *signals* can be represented by the elements of L_m . Accordingly, a CT system Δ with k -dimensional input v and q -dimensional output w is viewed as a subset $\Delta \subset L_k \times L_q$. Let us call such system $\Delta \subset L_k \times L_q$ *weakly causally stable* if for every $T \geq 0$, $(v, w) \in \Delta$, and $v_* \in L_k^2$ such that $v(t) = v_*(t)$ for all $t \leq T$ there exists a sequence of elements $(v_i, w_i) \in \Delta \cap (L_k^2 \times L_q^2)$, such that $v_i(t) = v(t)$ and $w_i(t) = w(t)$ for all $t \leq T$, and $\|v_i - v_*\| \rightarrow 0$ as $i \rightarrow \infty$.

Given real matrices $A \in \mathbb{R}^{n,n}$, $B_1 \in \mathbb{R}^{n,k}$, $B_2 \in \mathbb{R}^{n,q}$, $Q \in \mathbb{R}^{n+k+q,n+k+q}$, where A is a Hurwitz matrix, and a subset $X_0 \subset \mathbb{R}^n$, let us say that system $\Delta \subset \ell_k \times \ell_q$ satisfies the *conditional* Integral Quadratic Constraint (IQC) defined by A, B_1, B_2, Q, X_0 if there exists a continuous function $\kappa : \mathbb{R}^n \times \mathbb{R}^k \times \mathbb{R}^q \mapsto \mathbb{R}$ such that

$$\int_0^\infty \sigma(x(t), v(t), w(t)) dt \geq -\kappa(x_0, v(0), w(0)) \quad (2.30)$$

for all $(v, w) \in \Delta \cap (L_k^2 \times L_q^2)$, $x_0 \in X_0$, where $x(\cdot)$ is defined by $v(\cdot)$, $w(\cdot)$, and x_0 according to

$$\dot{x}(t) = Ax(t) + B_1v(t) + B_2w(t), \quad x(0) = x_0, \quad (2.31)$$

which is understood, in a generalized sense, as

$$x(t) = x_0 + \int_0^t [Ax(\tau) + B_1v(\tau) + B_2w(\tau)]d\tau \quad (t \geq 0).$$

Similarly, let us say that $\Delta \subset L_k \times L_q$ satisfies the *complete IQC* defined by A , B_1 , B_2 , Q , X_0 if there exists a continuous function $\kappa : \mathbb{R}^n \times \mathbb{R}^k \times \mathbb{R}^q \mapsto \mathbb{R}$ such that

$$\int_0^T \sigma(x(t), v(t), w(t))dt \geq -\kappa(x_0, v(0), w(0)) \quad (T \geq 0) \quad (2.32)$$

for all $(v, w) \in \Delta$, $x_0 \in X_0$, and $x \in L_n$ satisfying (2.31).

Theorem 2.4 *Let $\Delta \subset L_k \times L_q$ be a weakly causally stable system which satisfies the conditional IQC defined by real matrices $A \in \mathbb{R}^{n,n}$, $B_1 \in \mathbb{R}^{n,k}$, $B_2 \in \mathbb{R}^{n,q}$, $Q \in \mathbb{R}^{n+k+q, n+k+q}$, where A is a Hurwitz matrix, and a subset $X_0 \subset \mathbb{R}^n$. Assume that*

- (a) *condition (2.29), where Π_{ij} are defined by (2.19) and (1.4) with $B = [B_1, B_2]$, is satisfied for some $\epsilon > 0$ and $s_0 \in j\mathbb{R}$;*
- (b) *there exist real matrices $C \in \mathbb{R}^{k,n}$, $D_1 \in \mathbb{R}^{k,k}$, $D_2 \in \mathbb{R}^{k,q}$ such that the quadratic form σ defined by (2.18) satisfies the inequality (2.26), and*

$$\det \begin{bmatrix} A - sI & B_1 \\ C & D_1 \end{bmatrix} \neq 0 \quad \forall s \in \mathbb{C}_+. \quad (2.33)$$

Then Δ satisfies the complete IQC defined by A, B_1, B_2, Q, X_0 .

3 Appendix

This section contains proofs of main statements made in the paper, including a brief description of some classical mathematical constructions used in the proofs.

3.1 Quadratic Optimization and Minimax

We begin by summarizing some elementary statements concerning quadratic functionals and real Hilbert spaces.

3.1.1 Quadratic Forms

A function $\sigma : V \mapsto \mathbb{R}$ defined on a real vector space V is called a *quadratic form* when $\sigma(v) = b(v, v)$ for all $v \in V$, where $b : V \times V \mapsto \mathbb{R}$ is a *symmetric bilinear* function, i.e.

$$b(u, v) = b(v, u), \quad b(u, xv + yw) = xb(u, v) + yb(u, w) \quad \forall u, v, w \in V, x, y \in \mathbb{R}. \quad (3.34)$$

This correspondence between symmetric bilinear functions and quadratic forms is a bijection, as $b(\cdot, \cdot)$ can be recovered from $\sigma(\cdot)$ according to the identity

$$b(u, v) = \frac{\sigma(u + v) - \sigma(u - v)}{4}.$$

The quadratic form σ is called *positive definite* (notation $\sigma \gg 0$) when $\sigma(v) > 0$ for all $v \neq 0$, and *positive semidefinite* (notation $\sigma \geq 0$) when $\sigma(v) \geq 0$ for all $v \in V$. Due to the identity

$$t\sigma(v) + (1 - t)\sigma(u) - \sigma(tv + (1 - t)u) = t(1 - t)\sigma(v - u) \quad \forall v, u \in V, t \in \mathbb{R}, \quad (3.35)$$

which is valid for every quadratic form $\sigma : V \mapsto \mathbb{R}$, σ is convex if and only if it is positive semidefinite.

For example, a symmetric real matrix $Q = Q' \in \mathbb{R}^{n,n}$ defines a symmetric bilinear form $b_Q : \mathbb{R}^n \times \mathbb{R}^n \mapsto \mathbb{R}$ according to $b_Q(v, u) = v'Qu$, and the associated quadratic form $\sigma_Q(v) = b_Q(v, v)$; the form σ_Q (equivalently, the matrix $Q = Q'$) is positive definite (or semidefinite) when all eigenvalues of Q are positive (notation $Q > 0$) or, respectively, non-negative (notation $Q \geq 0$).

3.1.2 Quadratic Optimization and Real Hilbert Spaces

In this paper, the terminology of quadratic forms is used to *formulate* statements (this makes assumptions easier to verify in applications), while the more flexible Hilbert space viewpoint is employed in the corresponding *proofs*. The definitions and statements of this subsection facilitate easy switching between the two frameworks.

Let $b : V \times V \mapsto \mathbb{R}$ be a symmetric bilinear form on a real vector space V such that the corresponding quadratic form $\sigma(v) = b(v, v)$ is positive definite. Since the quadratic function

$$t \in \mathbb{R} \mapsto \sigma(v + tu) = \sigma(v) + 2tb(v, u) + t^2\sigma(u)$$

takes only non-negative values, its discriminant is not positive, which yields the Cauchy-Schwartz inequality

$$|b(v, u)|^2 \leq \sigma(v)\sigma(u) \quad \forall v, u \in V, \quad (3.36)$$

and in turn implies that the function $v \mapsto |v|_\sigma = \sigma(v)^{1/2}$ is a *norm* on V .

Let V^σ be the set of all linear functions $f : V \mapsto \mathbb{R}$ such that

$$|f|_\sigma \stackrel{\text{def}}{=} \sup\{f(v) : \sigma(v) \leq 1\} < \infty.$$

As a dual of a normed space $(V, |\cdot|_\sigma)$, the pair $(V^\sigma, |\cdot|_\sigma)$ defines a Banach space. Let $\pi_\sigma : V \mapsto V^\sigma$ be the "natural correspondence" mapping every $v \in V$ to $f = \pi_\sigma v \in V^\sigma$ according to $f(u) = b(v, u)$.

The quantity $|f|_\sigma^2$, where $f \in V^\#$ can also be interpreted as the minimal upper bound in the *quadratic optimization* task

$$2fv - \sigma(v) \mapsto \sup_{v \in V}, \quad (3.37)$$

because

$$\inf_v \{2f(v) - \sigma(v)\} = \inf_{\sigma(v) \leq 1} \inf_{t \in \mathbb{R}} \{2f(tv) - \sigma(tv)\} = \inf_{\sigma(v) \leq 1} \inf_{t \in \mathbb{R}} \{2f(v)t - \sigma(v)t^2\}.$$

Theorem 3.1 *Let $b : V \times V \mapsto \mathbb{R}$ be a symmetric bilinear form on a real vector space V such that the corresponding quadratic form $\sigma(v) = b(v, v)$ is positive definite. Then*

- (a) *the set $\pi_\sigma V$ is dense in $(V^\sigma, |\cdot|_\sigma)$;*
- (b) *there exists a (unique) symmetric bilinear form $\bar{b} : V^\sigma \times V^\sigma \mapsto \mathbb{R}$ such that $|f|_\sigma^2 = \bar{b}(f, f)$ for all $f \in V^\sigma$;*
- (c) *for the bilinear form \bar{b} defined in (b), the identity $f(v) = \bar{b}(f, \pi_\sigma v)$ holds for all $f \in V^\sigma$ and $v \in V$.*

Theorem 3.1 establishes $(V^\sigma, |\cdot|_\sigma)$ as a (real) Hilbert space, and provides a linear norm-preserving bijection π_σ between vectors from V and elements of a dense subspace $\pi_\sigma V$ of V^σ . It also shows that the minimal upper bound in quadratic optimization (3.37), as a function of $f \in V^\#$, is a quadratic form on the subset V^σ where its values are finite.

In this paper, we will use either $|w|$ or $|w|_H$ to denote the norm of a vector w in a Hilbert space H . In addition, the shortcut $v'u$ will denote the scalar product of two vectors v, u from the same Hilbert space H . This notation can be motivated by the natural association of vectors $v \in H$ with bounded linear functions $L_v : \mathbb{R} \mapsto H$ defined by $L_v(t) = tv$. Accordingly, the adjoint v' is the linear function $v' : H \mapsto \mathbb{R}$ mapping u to the scalar product of v and u , and the composition $v'u$, where $v, u \in H$, is a linear function mapping \mathbb{R} to \mathbb{R} , i.e. a real number, which equals the scalar product of v and u .

3.1.3 Quadratic Minimax

The following statement lists sufficient conditions for the minimax identity in quadratic optimization.

Theorem 3.2 *Let V, W be real vector spaces. Let $g : V \times W \mapsto \mathbb{R}$ be defined by*

$$g(v, w) = \sigma(v) + 2p(v, w) - \mu(w) - 2f(v) + 2h(w) + r, \quad (3.38)$$

where $\sigma : V \mapsto \mathbb{R}$, $\mu : W \mapsto \mathbb{R}$, $p : V \times W \mapsto \mathbb{R}$, $f : V \mapsto \mathbb{R}$, $h : W \mapsto \mathbb{R}$, and $r \in \mathbb{R}$ are two positive definite quadratic forms, a bilinear functional, two linear functions, and a real number. Assume that

- (i) there exists $c \geq 0$ such that $c^2\sigma(v)\mu(w) \geq |p(v, w)|^2$ for all $v \in V$, $w \in W$;
- (ii) $\inf_{v \in V} g(v, 0) > -\infty$ and $\sup_{w \in W} g(0, w) < +\infty$.

Then

- (a) the minimax equality (2.15) holds;
- (b) there exists a constant $c_1 \geq 0$ such that

$$\inf_{v \in V} g(v, w) \geq -c_1(1 + \mu(w)) \quad \forall w \in W, \quad \sup_w g(v, w) \leq c_1(1 + \sigma(v)) \quad \forall v \in V.$$

Proof. By (i), for every $v \in V$ the function $f_v : W \mapsto \mathbb{R}$ defined by $f_v(w) = p(v, w)$ is linear and satisfies $|f_v|_\mu \leq c|v|_\sigma$, i.e. $f_v \in W^\mu$. Since the corresponding function $\pi_\sigma v \in V^\sigma \mapsto f_v \in W^\mu$ is linear and bounded, it can be extended to a bounded linear operator $L : V^\sigma \mapsto W^\mu$ such that

$$p(v, w) = \bar{w}' L \bar{v} \quad (\bar{w} = \pi_\mu w, \bar{v} = \pi_\sigma v) \quad \forall v \in V, w \in W.$$

Since L is bounded, its adjoint L' is well defined and bounded as well. Also, by (ii), $f \in V^\sigma$ and $h \in W^\mu$, hence the identity $g(v, w) = \bar{g}(\bar{v}, \bar{w})$ holds for $\bar{w} = \pi_\mu w$, $\bar{v} = \pi_\sigma v$, and

$$\bar{g}(\bar{v}, \bar{w}) = |\bar{v}|^2 + 2\bar{w}' L \bar{v} - |\bar{w}|^2 - 2f' \bar{v} + 2h' \bar{w} + r.$$

Let $A : V^\sigma \times W^\mu \mapsto V^\sigma \times W^\mu$ be the linear operator with block representation

$$A = \begin{bmatrix} I & L' \\ -L & I \end{bmatrix},$$

i.e. $A(v, w) = (v + L'w, w - Lv)$. Since A is bounded and $A + A' = 2I$ is strictly positive definite, A must be invertible, and hence there exist $v_0 \in V^\sigma$, $w_0 \in W^\mu$ such that $A(v_0, w_0) = (f, h)$.

Since L is bounded and the subsets $\pi_\sigma V$, $\pi_\mu W$ are dense in V^σ and W^μ respectively, we have (using notation $\bar{v} = \pi_\sigma v$, $\bar{w} = \pi_\mu(w)$ for $v \in V$ and $w \in W$):

$$\begin{aligned} \sup_{w \in W} g(v, w) &= \sup_{\bar{w} \in \pi_\mu W} \bar{g}(\bar{v}, \bar{w}) \\ &= \sup_{\hat{w} \in W^\mu} \bar{g}(\bar{v}, \hat{w}) \\ &= |\bar{v}|^2 + |L\bar{v} + h|^2 - 2f'\bar{v} + r \\ &= |\bar{v}|^2 + |L\bar{v} - Lv_0 + w_0|^2 - 2(v_0 + L'w_0)'\bar{v} + r \\ &= |\bar{v} - v_0|^2 + |L(\bar{v} - v_0)|^2 + |w_0|^2 - |v_0|^2 - 2w_0'Lv_0 + r, \end{aligned}$$

hence the second inequality in (b) holds, and

$$\begin{aligned}\inf_{v \in V} \sup_{w \in W} g(v, w) &= \inf_{\bar{v} \in \pi_\sigma V} \sup_{\bar{w} \in \pi_\mu W} \bar{g}(\bar{v}, \bar{w}) \\ &= |w_0|^2 - |v_0|^2 - 2w'_0 L v_0 + r.\end{aligned}$$

Similarly,

$$\begin{aligned}\inf_{v \in V} g(v, w) &= \inf_{\bar{v} \in \pi_\sigma V} \bar{g}(\bar{v}, \bar{w}) \\ &= \inf_{\hat{v} \in V_\sigma} \bar{g}(\hat{v}, \bar{w}) \\ &= -|\bar{w}|^2 - |L'\bar{w} - f|^2 + 2h'\bar{w} + r \\ &= -|\bar{w}|^2 - |L'\bar{w} - L'w_0 - v_0|^2 + 2\bar{w}'(w_0 - Lv_0) + r \\ &= -|\bar{w} - w_0|^2 - |L'(\bar{w} - w_0)|^2 + |w_0|^2 - |v_0|^2 - 2w'_0 L v_0 + r,\end{aligned}$$

hence the first inequality in (b) holds, and

$$\sup_{w \in W} \inf_{v \in V} g(v, w) = \sup_{\bar{w} \in \pi_\mu W} \inf_{\bar{v} \in \pi_\sigma V} \bar{g}(\bar{v}, \bar{w}) = |w_0|^2 - |v_0|^2 - 2w'_0 L v_0 + r,$$

which establishes the minimax identity.

The bounds from (b) follow from the explicit expressions for the partial optimal values, and from the boundedness of L and L' . ■

3.2 KYP Lemma Proofs

This section contains proofs of the statements associated with the KYP Lemma.

3.2.1 Theorem 1.1, (a) \Rightarrow (b)

For $z \notin \Lambda(A)$ let

$$L(z) = (zI - A)^{-1}B, \quad H(z) = D + CL(z).$$

Substituting $x = L(z)u$ (which means $Ax + Bu = zx$) with $z \in \mathbb{T}$ into (1.2) yields

$$\Pi(z) = H(z)'H(z) \quad \forall z \in \mathbb{T} \setminus \Lambda(A), \quad (3.39)$$

hence $\Pi(z) \geq 0$ for $z \in \mathbb{T} \setminus \Lambda(A)$. Moreover, since

$$\det \begin{bmatrix} z^{-1}A - I & z^{-1}B \\ C & D \end{bmatrix} = \det H(z)$$

for $z \neq 0$, $z \notin \Lambda(A)$, the rational function $z \mapsto \det H(z)$ is not identically equal to zero, and hence $\det H(z) \neq 0$ for all $z \in \mathbb{C}$ except, possibly, a finite subset. Hence (3.39) implies that $\Pi(z)$ is positive definite for all $z \in \mathbb{C}$ except, possibly, a finite subset.

3.2.2 Theorem 1.1, (b) \Rightarrow (a)

To prove the implication, we consider the associated optimization setup (1.5)-(1.7), which can be recognized as a case of quadratic optimization. The key step is to show that the infimum in (1.5)-(1.7) is finite. Then, according to Theorem 3.1, $\inf \Phi$ is a quadratic form of a . We define $P = P'$ by $\inf \Phi = -a'Pa$, and use the Bellman equation from dynamic programming to show that conditions (1.2),(1.3) are satisfied.

- (a) Let ℓ_m^2 be the set of complex square summable sequences $w : \mathbb{Z}_+ \mapsto \mathbb{C}$, equipped with the natural structure of a *real* vector space. Since $A + BK$ is a Schur matrix, there is a linear one-to-one correspondence between the pairs (x, u) in (1.6),(1.7) and the pairs $(w, a) \in \ell_m^2 \times \mathbb{C}^n$ which maps (x, u) to $(u - Kx, x(0))$.

Using the Parseval identity, the functional Φ in (1.5) can be re-written in the form

$$\Phi = \int_{\mathbb{T}} \{ \hat{w}(z)' \alpha(z) \hat{w}(z) + 2\text{Re}[\hat{w}(z) \beta(z) a] + a' \gamma(z) a \} dm(z), \quad (3.40)$$

where

$$\int_{\mathbb{T}} f(z) dm(z) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(e^{j\theta}) d\theta$$

denotes the standard Lebesgue measure integral on the unit circle \mathbb{T} ,

$$\hat{w}(z) = \sum_{t=0}^{\infty} w(t) z^{-t}$$

is the Fourier transform of $w \in \ell_m^2$, a square integrable function $\hat{w} : \mathbb{T} \mapsto \mathbb{C}$, and α, β, γ are the rational matrix-valued functions defined by the block decomposition identity (to be satisfied for $z \in \mathbb{T}$)

$$\begin{bmatrix} \alpha(z) & \beta(z)' \\ \beta(z) & \gamma(z) \end{bmatrix} = M(z)' Q M(z),$$

with

$$M(z) = \begin{bmatrix} I & 0 \\ K & I \end{bmatrix} \begin{bmatrix} (zI - A - BK)^{-1} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} B & I \\ I & 0 \end{bmatrix}.$$

Since $A + BK$ is a Schur matrix, α, β, γ have no poles on the unit circle \mathbb{T} . Also, since

$$\alpha(z) = F(z)' \Pi(z) F(z), \quad \text{where } F(z) = [I - K(z - A)^{-1} B]^{-1}$$

for $z \in \mathbb{T}$, the matrix $\alpha(z)$ is positive definite for all $z \in \mathbb{T}$ except, possibly, a finite subset, where it is positive semidefinite.

Since, at the points where $\alpha(z)$ is positive definite,

$$\bar{w}' \alpha(z) \bar{w} + 2\text{Re} \bar{w}' \beta(z) a \geq -a' \beta(z)' \alpha(z)^{-1} \beta(z) a \quad \forall \bar{w} \in \mathbb{C},$$

the conclusion $\inf \Phi > -\infty$ can be reached easily when there exists a constant $c \in \mathbb{R}$ such that $\beta(z)' \alpha(z)^{-1} \beta(z) \leq c I_m$ for all $z \in \mathbb{T}$ with $\alpha(z) > 0$. While such $c \in \mathbb{R}$ does not always exist, we can use the fact that

$$\int_{z \in \mathbb{T}} \hat{w}(z)' \delta(z) dm(z) = 0$$

for every $w \in \ell_m^2$ and every strictly proper rational matrix $\delta = \delta(z)$ with no poles outside the open unit disk $|z| < 1$.

Indeed, to prove that $\inf \Phi > -\infty$, it is sufficient to find a strictly proper rational matrix function $\delta = \delta(z)$ with no poles outside the open unit circle $|z| < 1$, with the property that there exists a constant $c \in \mathbb{R}$ such that

$$(\beta(z) - \delta(z))' \alpha(z)^{-1} (\beta(z) - \delta(z)) \leq c I_m \quad \text{for } z \in \mathbb{T} : \alpha(z) > 0. \quad (3.41)$$

Let

$$R = \max_{z \in \mathbb{T}} \lambda_{\max}(\alpha(z))$$

be the maximal eigenvalue of $\alpha(z)$ over $z \in \mathbb{T}$ (it exists since α is continuous on \mathbb{T}). Then $\alpha(z) \geq \rho(z) I_m$ for all $z \in \mathbb{T}$, where the scalar rational function $\rho = \rho(z)$ is defined by

$$\rho(z) = \det(\alpha(z)) R^{1-m}.$$

Hence condition (3.41) will be satisfied, for some $c \in \mathbb{R}$, when the ratio $(\beta - \delta)/\rho$ is bounded on \mathbb{T} , i.e. when the unit circle zeroes of the scalar components of $\beta - \delta$ match (counting multiplicity) the unit circle zeroes of ρ .

Recall that for every set of distinct complex numbers $(\lambda_i)_{i=1}^N$ and polynomials

$$p_i(\lambda) = \sum_{l=0}^{m_i-1} p_{i,l} \lambda^l$$

there exists a polynomial $p = p(\lambda)$ of degree $\sum m_i$ such that

$$p(\lambda) - p_i(\lambda) = O((\lambda - \lambda_i)^{m_i}) \quad \text{as } \lambda \rightarrow \lambda_i \quad \forall i.$$

Hence the boundedness of $(\beta - \delta)/\rho$ on \mathbb{T} can be achieved by selecting $\delta = \delta(z)$ as a linear combination of a sufficiently large number of monomials z^{-i} with positive integer i , which completes the proof of the inequality $\inf \Phi > -\infty$.

- (b) Since $V(a) \stackrel{\text{def}}{=} \inf \Phi > -\infty$ for every $a \in \mathbb{C}^n$, Theorem 3.1, together with representation (3.40), imply that $V = V(a)$ is a quadratic form of $a \in \mathbb{C}^n$. Moreover, since multiplying a solution (x, u) of (1.6) with $x(0) = a$ by j yields a solution (jx, ju) of (1.6) with $x(0) = ja$ and the same value of Φ , we have $V(ja) = V(a)$ for every

$a \in \mathbb{C}^n$, which implies that $V(a) = -a'Pa$ for some fixed complex n -by- n matrix $P = P'$. The Bellman inequality for the optimization task (1.5)-(1.7) can be written in the form

$$\inf_{u \in \mathbb{C}^m} \{\sigma(x, u) + V(Ax + Bu) - V(x)\} = 0. \quad (3.42)$$

Since $\mu(x, u) \stackrel{\text{def}}{=} \sigma(x, u) + V(Ax + Bu) - V(x)$ is a quadratic form in (x, u) , condition (3.42) means that $\mu(x, u) = |Cx + Du|^2$ for some $C \in \mathbb{C}^{m,n}$ and $D \in \mathbb{C}^{m,m}$ such that D is not singular. In other words, representation (1.2) takes place, and the inequality in (1.3) is satisfied for $\lambda = 0$.

To show that the inequality in (1.3) is satisfied for $0 < |\lambda| < 1$, note that otherwise there exist $p \in \mathbb{C}^n$, $q \in \mathbb{C}^m$, and $\xi \in \mathbb{C}$ such that

$$\begin{bmatrix} p \\ q \end{bmatrix}' \begin{bmatrix} \xi A - I & \xi B \\ C & D \end{bmatrix} = 0, \quad \begin{bmatrix} p \\ q \end{bmatrix} \neq 0, \quad |\xi| \in (0, 1).$$

Then $q \neq 0$ (otherwise $\xi p'A = p$, $p'B = 0$, $p \neq 0$ and hence the pair (A, B) is not stabilizable), and therefore it is possible to re-scale (p, q) in such a way that $|q| = 1$. Hence, for a solution x, u of (1.5)

$$\begin{aligned} |Cx(t) + Du(t)|^2 &\geq |q'Cx(t) + q'Du(t)|^2 \\ &= |p'x(t) - \xi p'Ax(t) - \xi p'Bu(t)|^2 \\ &= |p'x(t) - \xi p'x(t+1)|^2, \end{aligned}$$

which implies that

$$\sum_{t=0}^{\infty} |Cx(t) + Du(t)|^2 \geq (1 - |\xi|^2) |p'a|^2, \quad (3.43)$$

contradicting the construction of C, D , which guarantees that the maximal lower bound of the left side in (3.43) is zero for all $a \in \mathbb{C}^n$.

3.2.3 Theorem 1.1, the Case of Real Coefficients

When the matrices A, B, Q in (b) are real, for every solution (x, u) of (1.6) with $x(0) = a$ the conjugated pair (\bar{x}, \bar{u}) is a solution of (1.6) with $x(0) = \bar{a}$ and the same value of Φ . Hence $V(\bar{a}) = V(a)$ for every $a \in \mathbb{C}^n$, which implies that the (generally complex) matrix $P = P'$ in the representation $V(a) = -a'Pa$ actually has real coefficients. Since, in this case, the Hermitian form $\sigma(x, u) - V(x) + V(Ax + Bu)$ has real coefficients, the matrices C, D can also be chosen to be real.

3.2.4 Proof of Theorem 1.2

By (1.2) we have

$$\Phi = -a'Pa + \sum_{t=0}^{\infty} |Cx(t) + Du(t)|^2,$$

and it was already shown in the proof of Theorem 1.1, (b) \Rightarrow (a) that $\inf \Phi = -a'Pa$. Hence Φ converges to its maximal lower bound if and only if the sum of squares of $Cx + Du$ converges to zero.

3.2.5 Proof of Theorem 1.3

The implication (a) \Rightarrow (b) is trivial, as substituting a non-zero pair (x, u) from $\mathcal{L}(z)$ with $|z| = 1$ into (1.8) yields $\sigma(x, u) = \sigma_P(x, u) > 0$.

To prove that (b) implies (a), assume that (b) is true but (a) is not, which means that 0 is not in the convex set

$$\Omega = \{Q + E'_0PE_0 - E'_1PE_1 - S : S = S' > 0, P = P'\},$$

where

$$E_0 = [I_n \ 0], \quad E_1 = [A \ B].$$

According to the Hahn-Banach theorem there exists a hyperplane which separates (non-strictly) Ω from zero, i.e. there exists matrix $H = H' \neq 0$ such that

$$\text{tr}(XH) \leq 0 \quad \forall X \in \Omega. \quad (3.44)$$

Using (3.44) with $X = Q - tI$ where $t \rightarrow 0$ yields $\text{tr}(QH) \leq 0$. Using (3.44) with $X = Q - I - tpp'$ where $t \rightarrow +\infty$ yields $\text{tr}(Hpp') \geq 0$ for every $p \in \mathbb{C}^{n+m}$, i.e. $H \geq 0$. Similarly, using (3.44) with $X = Q - I + t(E'_0PE_0 - E'_1PE_1)$ where $t \rightarrow \pm\infty$ yields $\text{tr}(H(E'_0PE_0 - E'_1PE_1)) = 0$ for every $P = P'$, i.e. $E_0HE'_0 = E_1HE'_1$. The last equality implies existence of a unitary matrix U such that $UH^{1/2}E'_0 = H^{1/2}E'_1$, or, equivalently, $E_0H^{1/2}U' = E_1H^{1/2}$. Let w_1, \dots, w_{n+m} be an orthonormal basis of eigenvectors of U' , with $z_i \in \mathbb{T}$ being the corresponding eigenvalues. Define $x_i \in \mathbb{C}^n$, $u_i \in \mathbb{C}^m$ by

$$e_i = \begin{bmatrix} x_i \\ u_i \end{bmatrix} = H^{1/2}w_i.$$

By construction, $(x_i, u_i) \in \mathcal{L}(z_i)$, and hence by assumption (b) $e'_iQe_i > 0$ whenever $e_i \neq 0$. On the other hand

$$0 \geq \text{tr}(QH) = \text{tr}\left(Q \sum_{i=1}^{n+m} e_i e'_i\right) = \sum_{i=1}^{n+m} e'_i Q e_i,$$

which means $e'_i Q e_i = 0$ for all i . Hence $e_i = 0$ for all i and therefore $H = 0$, which contradicts the construction.

To complete the proof, consider the case when A, B, Q have real coefficients. Then for every $P = P'$ such that $\sigma_P > 0$ we also have $\sigma_{\bar{P}} > 0$, and hence, for $\tilde{P} = 0.5(P + \bar{P})$,

$$\sigma_{\tilde{P}} = 0.5(\sigma_P + \sigma_{\bar{P}}) > 0.$$

3.2.6 Proof of Theorem 1.4

The implication (a) \Rightarrow (b) follows in the standard way by substituting an arbitrary pair (x, u) from $\mathcal{L}(z)$ with $|z| = 1$ into (1.8), which yields $\sigma(x, u) = \sigma_P(x, u) \geq 0$.

To prove that (b) implies (a), consider the optimization task (1.5)-(1.7), take any K such that $A + BK$ is a Schur matrix, and consider the Fourier transform representation of Φ given by (3.40). Since $\alpha(z) \geq 0$ for all $z \in \mathbb{T}$, we have $\inf \Phi > -\infty$ for $a = 0$. Therefore $\inf \Phi > -\infty$ for every $a \in \mathbb{C}^n$ which is reachable from $x(0) = 0$ in system (1.5). Since the pair (A, B) is controllable, we conclude that $\inf \Phi > -\infty$ for every $a \in \mathbb{C}^n$.

We now use the same arguments as in the proof of Theorem 1.1 to establish that $\inf \Phi = -a'Pa$ for some matrix $P = P'$ (real whenever A, B, Q are real). Finally, positive semidefiniteness of σ_P follows from the Bellman equation.

3.2.7 KYP Proofs in Continuous Time

In principle it is possible to translate all steps in the proofs of Theorems 1.1-1.4 into a continuous-time format. However, there is a simple way of deriving the CT versions from the DT ones.

Choose $r > 0$ in such a way that the matrix $rI - A$ is not singular. Let $\bar{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$. Consider the bijection $h_0 : \bar{\mathbb{C}} \mapsto \bar{\mathbb{C}}$ and the linear bijection $h_1 : \mathbb{C}^n \times \mathbb{C}^m \mapsto \mathbb{C}^n \times \mathbb{C}^m$ which map $s \in \bar{\mathbb{C}}$ to $z = h_0(s) \in \bar{\mathbb{C}}$ and $(x, u) \in \mathbb{C}^n \times \mathbb{C}^m$ to $(\tilde{x}, u) = h_1(x, u) \in \mathbb{C}^n \times \mathbb{C}^m$ according to

$$\tilde{x} = \frac{rx - Ax - Bu}{\sqrt{2r}}, \quad z = \begin{cases} \infty, & s = r, \\ -1, & s = \infty, \\ \frac{r+s}{r-s}, & \text{otherwise.} \end{cases}$$

Define $\tilde{A}, \tilde{B}, \tilde{\sigma}$ by

$$\tilde{A} = (rI + A)(rI - A)^{-1}, \quad \tilde{B} = \sqrt{2r}(rI - A)^{-1}B,$$

and

$$\tilde{\sigma}(\tilde{x}, u) = \sigma(x, u) \quad \text{for } (\tilde{x}, u) = h_1(x, u)$$

(note that $\tilde{A}, \tilde{B}, \tilde{\sigma}$ will have real coefficients whenever A, B, σ have real coefficients).

Simple algebraic manipulations can be used to show that

- (a) for $(\tilde{x}, u) = h_1(x, u)$, equality $sx = Ax + Bu$ is satisfied if and only if $z\tilde{x} = \tilde{A}\tilde{x} + \tilde{B}u$ (including the case $s = \infty, z = -1$, in which case $sx = Ax + Bu$ is interpreted as $x = 0$, as well as the case $z = \infty, s = r$, in which case $z\tilde{x} = \tilde{A}\tilde{x} + \tilde{B}u$ is interpreted as $\tilde{x} = 0$);

(b) for every $(\tilde{x}, u) = h_1(x, u)$, the identity

$$2x'P(Ax + Bu) = (\tilde{A}\tilde{x} + \tilde{B}u)'P(\tilde{A}\tilde{x} + \tilde{B}u) - \tilde{x}'P\tilde{x}$$

holds;

(c) $h_0(j\mathbb{R} \cup \{\infty\}) = \mathbb{T}$;

(d) s is an eigenvalue of A if and only if $z = h_0(s)$ is an eigenvalue of \tilde{A} .

In order to prove the CT statements for some A, B, σ , choose r and construct $\tilde{A}, \tilde{B}, \tilde{\sigma}$ first. For $z \in \mathbb{T} \setminus \Lambda(\tilde{A})$ define $\tilde{\Pi} = \tilde{\Pi}(z)$ by the identity

$$u'\Pi(z)u = \tilde{\sigma}(\tilde{x}, u), \quad \text{subject to } z\tilde{x} = \tilde{A}\tilde{x} + \tilde{B}u.$$

Then $\Pi(s) = \tilde{\Pi}(z)$ for $z = h_0(s)$, i.e. the positive definiteness/semidefiniteness of $\tilde{\Pi}(z)$ on \mathbb{T} is determined by positive definiteness/semidefiniteness of $\Pi(s)$ for $s \in \mathbb{R} \cup \{\infty\}$. When matrices C, D are given, define \tilde{C}, \tilde{D} by the identity

$$Cx + Du = \tilde{C}\tilde{x} + \tilde{D}u \quad \text{for } (\tilde{x}, u) = h_1(x, u).$$

Now the DT statements of the KYP Lemma applied to $\tilde{A}, \tilde{B}, \tilde{\sigma}, \tilde{\Pi}$ (and, possibly, $\tilde{C}, \tilde{D}, \tilde{P} = P$) prove the corresponding CT statements of the KYP Lemma.

3.3 Minimax Theorem Proofs

This section contains the proofs of the minimax theorems associated with the KYP setup, as well as the corresponding IQC statements.

3.3.1 Proof of Theorems 2.1 and 2.3

The proof is based on associating the statements with the more general setup of Theorem 3.2.

In the DT case, let $V = \ell_k^2$, $W = \ell_q^2$. The functional g defined by (2.17),(2.18) is a quadratic form of $(v, w, a) \in V \times W \times \mathbb{R}^n$. Hence for every fixed $a \in \mathbb{R}^n$ it defines it defines unique quadratic forms σ, μ , bilinear form p , linear functions f, h , and a constant r such that representation (3.38) takes place. According to Theorem 1.1, condition $\Pi_{11}(z) \geq 0$ (for $z \in \mathbb{T}$), coupled with $\Pi_{11}(z_0) > 0$ (both parts of assumption (2.21)), implies that $g(v, 0)$ has a finite lower bound. Similarly, $\Pi_{22}(z) \leq 0$ (for $z \in \mathbb{T}$), coupled with $\Pi_{11}(z_0) < 0$, implies that $g(0, w)$ has a finite upper bound, so condition (ii) of Theorem 3.2 is satisfied. Finally, in terms of Fourier transforms we have

$$\sigma(v) = \int_{\mathbb{T}} \hat{v}' \Pi_{11} \hat{v} dm(z), \quad \mu(w) = \int_{\mathbb{T}} \hat{w}' \Pi_{22} \hat{w} dm(z), \quad p(v, w) = \operatorname{Re} \int_{\mathbb{T}} \hat{v}' \Pi_{12} \hat{w} dm(z).$$

Since (2.21) implies that the matrix

$$\int_{\mathbb{T}} \begin{bmatrix} \hat{v} & 0 \\ 0 & \hat{w} \end{bmatrix}' \begin{bmatrix} \Pi_{11} & \epsilon \Pi_{12} \\ \epsilon \Pi_{21} & -\Pi_{22} \end{bmatrix} \begin{bmatrix} \hat{v} & 0 \\ 0 & \hat{w} \end{bmatrix} dm(z) = \begin{bmatrix} \int_{\mathbb{T}} \hat{v}' \Pi_{11} \hat{v} dm(z) & \epsilon \int_{\mathbb{T}} \hat{v}' \Pi_{12} \hat{w} dm(z) \\ \epsilon \int_{\mathbb{T}} \hat{w}' \Pi_{21} \hat{v} dm(z) & - \int_{\mathbb{T}} \hat{w}' \Pi_{22} \hat{w} dm(z) \end{bmatrix}$$

is positive semidefinite for all $v \in \ell_k^2$, $w \in \ell_q^2$, condition (i) is satisfied with $c = \epsilon^{-1}$.

According to Theorem 3.2 this means that the minimax equality holds. The bounds in (b) are now established as well. Since the partial optimal values are quadratic functionals, the bounds establish their continuity.

The proofs for the CT case follow the same pattern, with CT Fourier transform replacing the DT version.

3.3.2 Proof of Theorems 2.2 and 2.4

Consider the DT case first. For every $(v_0, w_0) \in \Delta$, $x_0 \in X_0$, and $T > 0$ consider functional g from Theorem 2.1 defined with $a = x(T+1)$. According to (b), for every fixed $w \in \ell_q^2$ the maximal lower bound of $g(v, w)$ with respect $v \in \ell_k^2$ is not positive, i.e.

$$\sup_w \inf_v g(v, w) \leq 0.$$

By Theorem 2.1,

$$\inf_v \sup_w g(v, w) = \sup_w \inf_v g(v, w) \leq 0,$$

which means that there exists a sequence of signals $\{\tilde{v}_i\}_{i=1}^\infty \subset \ell_k^2$ such that $g(\tilde{v}_i, w) < 1/i$ for all $i \in \{1, 2, \dots\}$ and all $w \in \ell_q^2$. In addition, Theorem 2.1 also claims that $\sup_w g(v, w)$ is a continuous in the metric of the Hilbert space defined by the associated quadratic form σ . Since, for $v \in \ell_k^2$,

$$\sigma(v) = \int_{\mathbb{T}} \hat{v}(z)' \Pi_{11}(z) v(z) dm(z),$$

and Π_{11} is uniformly bounded on \mathbb{T} , the norm $\sigma(v)^{1/2}$ is majorated by the standard Hilbert space norm of ℓ_k^2 , and hence $\sup_w g(v, w)$ is a continuous in the standard metric of ℓ_k^2 . Accordingly, there exist a sequence of signals $\{\tilde{v}_i\}_{i=1}^\infty \subset \ell_k^2$ and a sequence of positive numbers $\delta_i > 0$ such that $g(u, w) < 1/i$ for all $i \in \{1, 2, \dots\}$, $u \in \ell_k^2$, and $w \in \ell_q^2$ such that $|u - \tilde{v}_i| < \delta_i$.

For every $i \in \{1, 2, \dots\}$ let

$$v_*(t) = \begin{cases} v(t), & t \leq T, \\ \tilde{v}_i(t - T - 1), & t > T. \end{cases}$$

Since Δ is assumed to be weakly causally stable, there exist $(\tilde{v}, \tilde{w}) \in \Delta \cap (\ell_k^2 \times \ell_q^2)$ such that $\tilde{v}(t) = v(t)$, $\tilde{w}(t) = w(t)$ for $t \leq T$, and $|\tilde{v} - v_*| < \delta_i$. Due to the way in which \tilde{v}_i, δ_i were chosen, for the corresponding solution \tilde{x} of

$$\tilde{x}(t+1) = A\tilde{x}(t) + B_1\tilde{v}(t) + B_2\tilde{w}(t), \quad \tilde{x}(0) = x_0$$

we have

$$\sum_{t>T} \sigma(\tilde{x}(t), \tilde{v}(t), \tilde{w}(t)) < \frac{1}{i}.$$

Since

$$\sum_{t=0}^{\infty} \sigma(\tilde{x}(t), \tilde{v}(t), \tilde{w}(t)) \geq \kappa(x_0, v(0), w(0))$$

by the conditional IQC assumption, and $x = \tilde{x}$, $v = \tilde{v}$, $w = \tilde{w}$ for $t \leq T$ for all i , we conclude (by letting $i \rightarrow \infty$) that

$$\sum_{t=0}^T \sigma(x(t), v(t), w(t)) \geq \kappa(x_0, v(0), w(0)),$$

which proves the complete IQC.

The derivation in the CT time case follows the same steps, with the definitions of a and v_* being modified to $a = x(T)$ and

$$v_*(t) = \begin{cases} v(t), & t \leq T, \\ \tilde{v}_i(t - T), & t > T. \end{cases}$$

4 Acknowledgements

This paper was written as a technical addendum to the Control Handbook article "Integral Quadratic Constraints" by Alexandre Megretski, Ulf T. Jönsson, Chung-Yao Kao, and Anders Rantzer. In particular, it establishes the claims of Theorem 2.1, and expands on the "minimax" approach to post-feasibility analysis mentioned there.

The author is very grateful to Ulf Jönsson and Anders Rantzer for their encouragement and many helpful comments.

References

- [1] Yakubovich, V.A. Solution of some matrix inequalities, met with in the theory of automatic control *Doklady Akademii Nauk SSSR*, v. 143, No. 6, p. 1304-1307, 21 April 1962.
- [2] A. L. Lihtarnikov, V. A. Jakubovich, The Frequency Theorem for Continuous One-Parameter Semigroups, *Mathematics of the USSR-Izvestiya* Volume 11, No. 4, p.849, 1977.
- [3] V. A. Yakubovich, Linear-quadratic optimization problem and the frequency theorem for periodic systems. II *Siberian Mathematical Journal*, Vol. 31, No. 6, p.1027-1039, 1980.

- [4] A. Rantzer, On the Kalman-Yakubovich-Popov Lemma *Systems and Control Letters*, 27:5, January 1996.